

הלק גאומטרי (1)

Convex optimization 1.1

(1) אלו $1 \leq i \leq m$ משפטים f_i הם פונקציות קמורות. נניח:

אם $\lambda_i = 0$ סתם. $\lambda_i > 0$ אז $0 \leq t \leq 1$

$$t f_i(x_1) + (1-t) f_i(x_2) \geq f_i(t x_1 + (1-t) x_2)$$

(נמצא פונקציה קמורה)

\Downarrow

$$\lambda_i (t f_i(x_1) + (1-t) f_i(x_2)) \geq \lambda_i f_i(t x_1 + (1-t) x_2)$$

\Downarrow

$$t \lambda_i f_i(x_1) + (1-t) \lambda_i f_i(x_2) \geq \lambda_i f_i(t x_1 + (1-t) x_2)$$

$\sum_{i=1}^m \lambda_i f_i$ פונקציה קמורה. סך המשפטים $\lambda_i f_i$

אז $0 \leq t \leq 1$

$$\sum_{i=1}^m t \lambda_i f_i(x_1) + \sum_{i=1}^m (1-t) \lambda_i f_i(x_2) = \sum_{i=1}^m t \lambda_i f_i(x_1) + (1-t) \sum_{i=1}^m \lambda_i f_i(x_2)$$

$$\geq \sum_{i=1}^m \lambda_i f_i(t x_1 + (1-t) x_2)$$

המשפט
 $\lambda_i f_i$



כל $f(x) = -x$

$g(x) = x^2$

(2) נניח:

אז $h(x) = f \circ g(x) = -x^2$

הפונקציה f קמורה, g קמורה
קמורה.



f קמורה: משפטים $0 \leq t \leq 1$

$$\left. \begin{aligned} t f(x_1) + (1-t) f(x_2) &= -t x_1 - (1-t) x_2 \\ f(t x_1 + (1-t) x_2) &= -t x_1 - (1-t) x_2 \end{aligned} \right\} \rightarrow t f(x_1) + (1-t) f(x_2) \geq f(t x_1 + (1-t) x_2)$$

g קמורה: הולכת בקושי 2.

h אינה קמורה. נניח הולכת בקושי 2

Sub-Gradient for Soft-Sum objective 1.2

$$f(w, b) = \max(0, 1 - y(x^T w + b)) \quad (3) \quad \text{לפי סוגריה } f \text{ רגילה:}$$

אם $z(w, b) = 0$, $\phi(w, b) = 1 - y(x^T w + b)$. אז f מתאגר מקסימום באינסוף ופ. /

עם זאת f סוקריה/לילר ורגל n כל סוקריה לילילר וילא קומה (גסרה 1):

אילא f סוקרי לילילר ורגל n , אז $f(x) = cx$ כלל $c \in \mathbb{R}$

$$f(tx + (1-t)x_2) = c(tx_1 + (1-t)x_2) = tx_1 + c(1-t)x_2 \quad \text{על } 0 \leq t \leq 1$$

$$= f(tx_1) + f((1-t)x_2) \quad \text{רגל.}$$

כל, הן הסיקריה ופ לילילר מרגל 1 ופ קילילר.

עם הוירג סוקרי כלל קוסל n מקסימום 0 סוקריה לילילר 0 וילא קומה

(הילא 9) ופ מקריה $f(w, b)$ וילא סוקריה קומה רגל.

$$g = \begin{cases} 0 \\ (-yx_1, -y) \end{cases}$$

$$l_{x,y}^{\text{hinge}}(w, b) = 0$$

(4)

$\mathbb{R}^d \ni y$ is given $1 \leq i \leq m$ is Subgradient \rightarrow 2.1.1 (5)

$$f_i(y) \geq f_i(x) + \langle g_i, y-x \rangle$$

\Downarrow

$$\sum_{i=1}^m f_i(y) \geq \sum_{i=1}^m f_i(x) + \langle g_i, y-x \rangle$$

$1 \leq k \leq m$ is given $1 \leq i \leq k$ is Subgradient \rightarrow 2.1.1

$$\sum_{i=1}^k g_i \in \partial \sum_{i=1}^k f_i(x) \quad \text{is given}$$



$$g_i = \begin{cases} 0 \\ (-y_i x_i, -y_i) \end{cases}$$

$$l_{x_i, y_i}^{hinge}(w, b) = \begin{cases} 0 \\ \text{else} \end{cases} \quad g_i$$

is given \rightarrow 2.1.1 (6)

$$\underline{\frac{1}{m} \sum_{i=1}^m g_i = \lambda(w, 0) \in \partial \frac{1}{m} \sum_{i=1}^m l_{x_i, y_i}^{hinge}(w, b) + \frac{\lambda}{2} \|w\|^2}$$

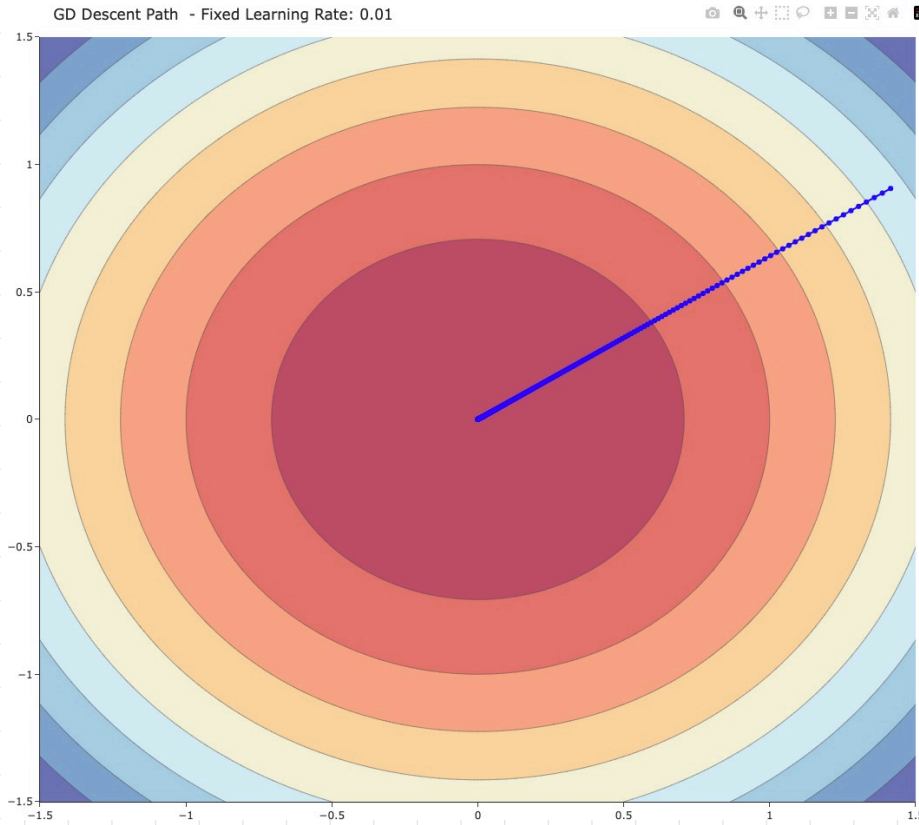


is given

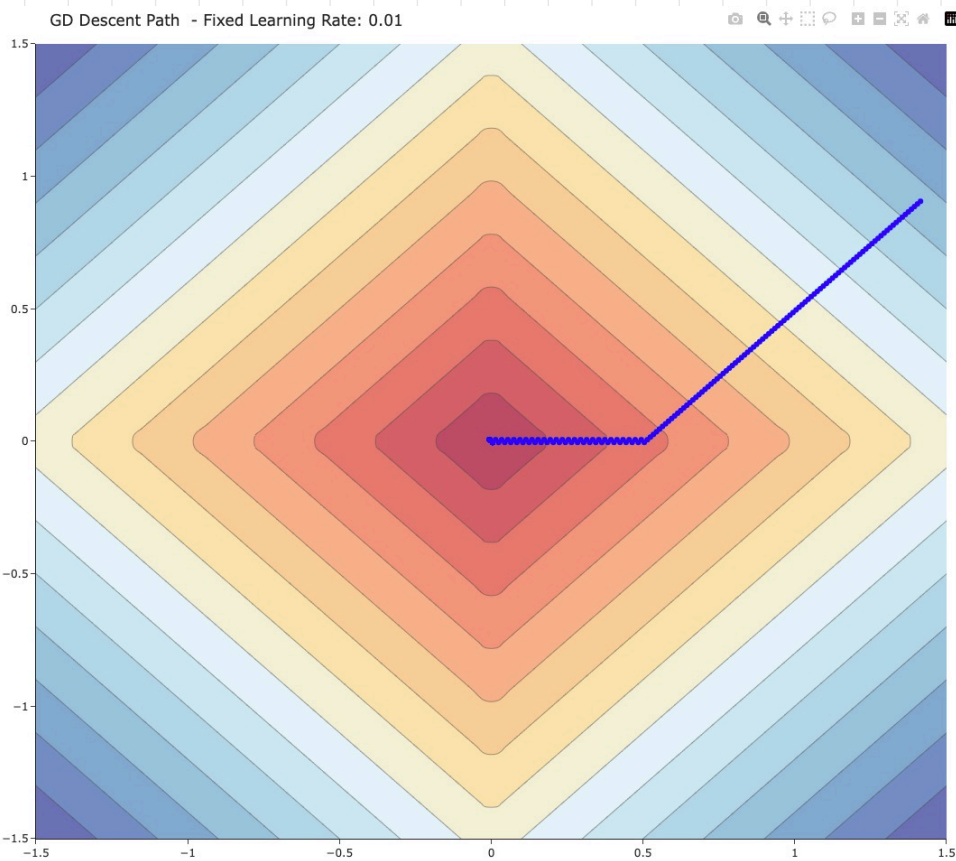
② למידה קבועה

Comparing fixed learning rates 2.1

①



למידה קבועה
 L_2



למידה קבועה
 L_1

והקב"ל ד'תתק"ל היה יפה אל התנאים שכן בני המלך ואלו מהחפץ בן המלך.
מורא בל משה בנעמו בירושלם. מרמז — ושימש בענייניו מרמז
ה וקורא בה שם מהתנאים והאצורה אהסין, חסן ואלוהים יאן
"לה" אל התנאים א י י ו דס.

בנעל אן. מורא ואל לשי שמש בנעמו ו מקיים כי ה בנקודת
בה שם יפה באברהם במחיק צורה מהתנאים, און האלוהים
אל יאן בצורה שיה אלא יאלל עמיתו וי יאן ע י י א יעני כיון
כמו שמקלען דמאנע וטניה.

② שמש זכ כי ומוא ול עשר מין זילגל בנעמו היטה אל התנאים
(נמשך אלא קלס היחיד כי יאן של גרמז מוינס ג'זילגל שמה אן
ועל. אן מנע שסדבז צו מתישר סדק טיגט מרמז ג'סין דמור
ול).

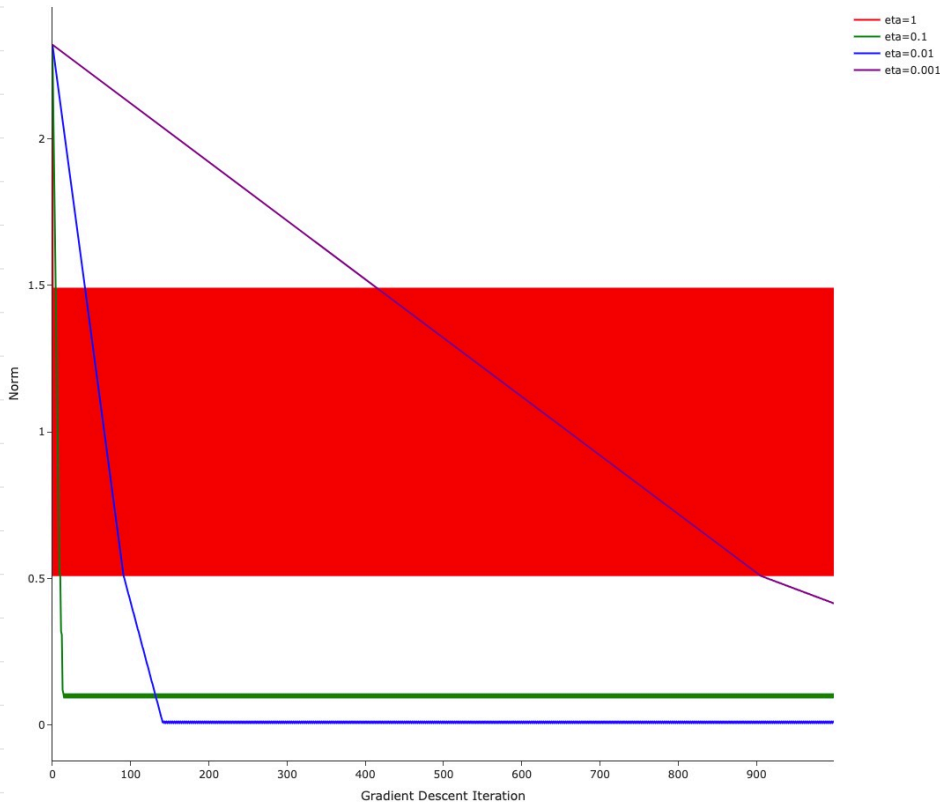
מקום. האלוהים במורא ול קודם יחד אל האלמסן א ג'המ?
כו ומו ממש סת' התנאים ויק' אלא מכן שלע כיון א'אלן טל
התנאים קצורה יסיון (ולו אלא ו להיחיד...).

④ ו דססל הי' תנאים מה שמש באר ול ומו צ-עס דקיון ואלה מל בל ומו צילס
בקיון קיים ובה סק בן השני, נסכי מור.

סל מור בל בעקמו מן סיוט שמש בנעמו ב ארמיו גל מלמיו ע'מ
א וויג' יחד און, קנקל ל גל ממש בנעמו לאיע שער. דה יאן
ג'סין דיון אלא מורא ול א מל מור בל כ'י שמזן דססל הי' תנאים
ג'סין דה אלא מרמז.

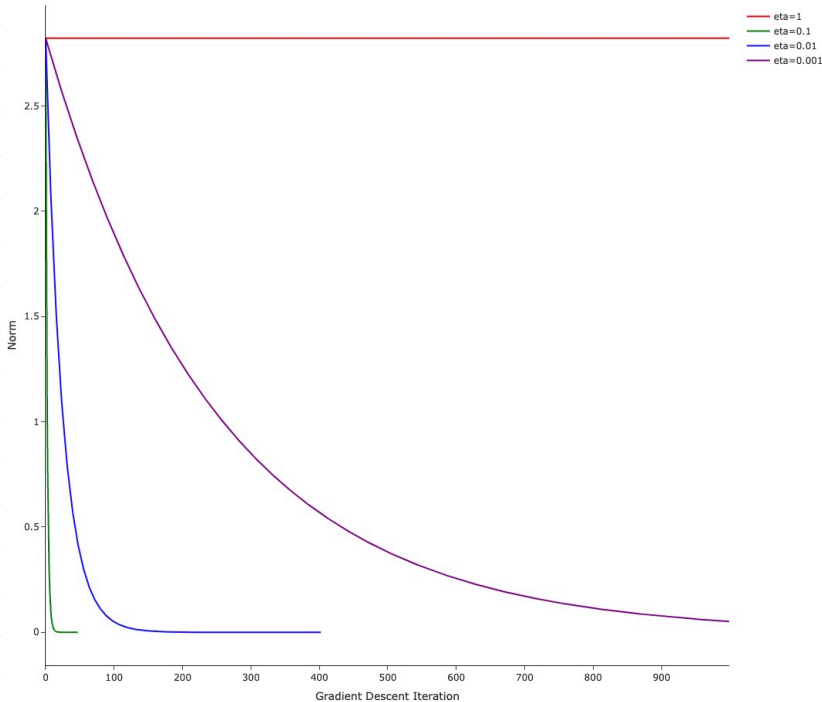
3

Convergence of L1

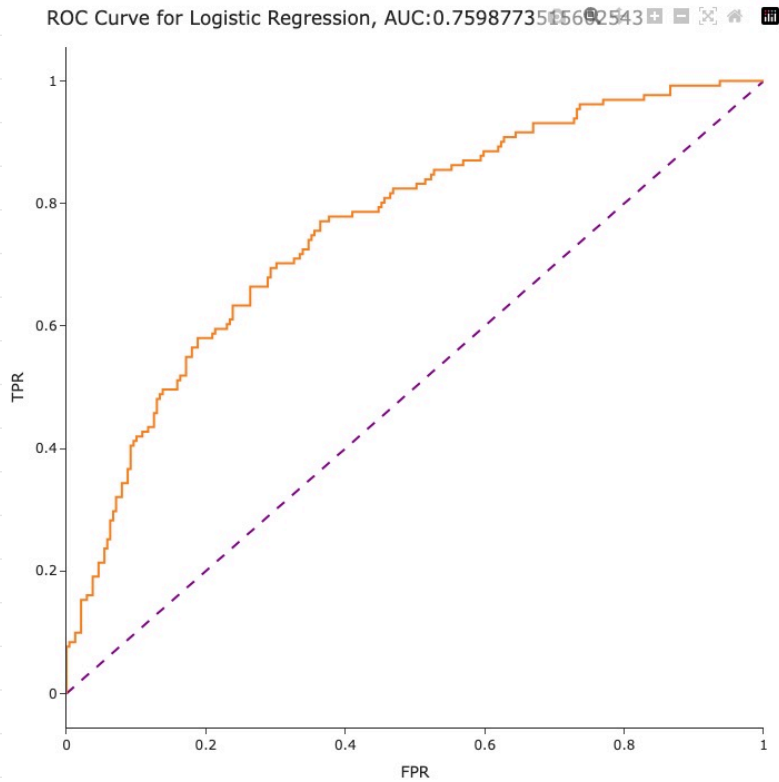


בגלל נשקף מהירותי. $\eta = 1$
 ה' חלש ונלוינו "למקדנר" ויז'ה ממקור
 במקדנר חלש לזוהי ג'ר, דג'ה וז'ה
 ילר ה"ט'קן וה'ק'ן ב'ק'ל.
 ד' ז'לר ה' חלש והמקדנר ג'ר'ל
 מ'ל'ר ד'ומה, ג'ל'ר ונ'ל'ר. ז'ר'ר
 ב'צ'ור א'ק'ר ונ'ל'ר ג'ר'ר א'ס,
 נ'ל'ר ב' ו $\eta = 0$ חלש ויז'ה
 ג'ל'ר וה'ר'ה ב'ל'ר.

Convergence of L2



כ'ו ב'ל'ר וה'ק'ר, א'ל'ר וה'א'ר
 ג'ט'ר ה' $\eta = 1$ חלש והמקדנר
 "ד'ו, מ'ל'ר מ'ל'ר ב'צ'ור מ'ל'ר
 ח'ס וה'ר'ר. א'ל'ר ונ'ל'ר וז'ל'ר.
 וה'ל'ר ונ'ל'ר חלש והמקדנר
 והמקדנר ה'א'ר ונ'ל'ר
 ויז'ה והמקדנר א'ס, נ'ל'ר
 נ'ל'ר מ'ל'ר ק'ל'ר וה'ל'ר
 א'ל'ר ו $\eta = 0$ חלש א'ל'ר
 מ'ק'ר מ'ל'ר מ'ל'ר וה'ר'ר
 וה'ר'ר ב'ל'ר.



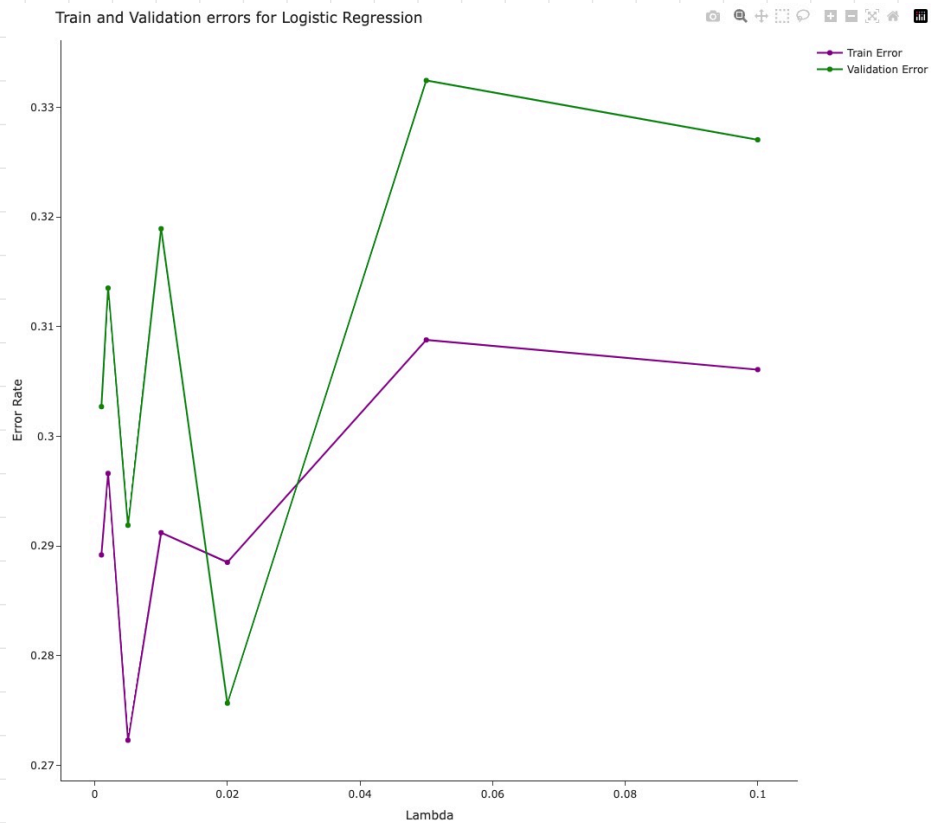
(5)

`argmax(TPR-FPR): 0.32477315611476026, test error: 0.33695652173913043`

(6)

הן נראו test error הן 0.3247/הן α , נראו
 . 0.3369 הן

Optimal lambda for l1: penalty: 0.02, Train Error: 0.322, Test Error: 0.2826086956521739



(כאמור, וזוהי האלווטיבליזציה של L_1 היא 0.02 / test error
 הוא 0.2826 בקירוב).